

ZHILIANG KONGZHI GUOCHENG  
ZHONG DE TONGJI JISHU

# 质量控制过程中 的统计技术

+++++  
+++++  
+++++

杨鑫 刘文长 主编  
张绍周 主审



化学工业出版社

ZHILIANG KONGZHI GUOCHENG  
ZHONG DE TONGJI JISHU

# 质量控制过程中 的统计技术

++++++  
++++++  
++++++

杨鑫 刘文长 主编  
张绍周 主审



化学工业出版社

·北京·

本书内容共分十章：①统计技术基本知识；②试验误差与数据处理；③从样本推断批产品质量示例；④过程控制中的统计技术；⑤质量管理统计工具；⑥显著性检验；⑦方差分析与试验设计；⑧回归分析；⑨统计抽样检验；⑩计算机在统计工作中的应用。读者可以根据工作需要，阅读其中有关章节。

本书适用于产品制造企业工程技术人员、质量检验人员、质量管理人員和质量统计人員阅读使用，也可供各级产品质量监督检验机构技术人员参考，是适用于各种产品制造企业检验人員和质量管理、统计人員的通用参考书籍。高校学生学习数理统计课程时也可参阅。

### 图书在版编目 (CIP) 数据

质量控制过程中的统计技术/杨鑫，刘文长主编.

—北京：化学工业出版社，2014.2

ISBN 978-7-122-19484-8

I. ①质… II. ①杨… ②刘… III. ①质量控制-统计方法 IV. ①F273.2

中国版本图书馆 CIP 数据核字 (2014) 第 002578 号

---

责任编辑：吕佳丽  
责任校对：边涛

文字编辑：郑直  
装帧设计：史利平

---

出版发行：化学工业出版社（北京市东城区青年湖南街13号 邮政编码100011）

印装：大厂聚鑫印刷有限责任公司

787mm×1092mm 1/16 印张17 字数427千字 2014年5月北京第1版第1次印刷

---

购书咨询：010-64518888（传真：010-64519686） 售后服务：010-64518899

网 址：<http://www.cip.com.cn>

凡购买本书，如有缺损质量问题，本社销售中心负责调换。

---

定 价：58.00 元

版权所有 违者必究

## 本书编写人员

主 编	杨 鑫	刘文长		
副 主 编	张丽军	赵艳茹	冉维民	张凤元
参编人员	刘 莉	戴 平	于克孝	覃爱平
	张瑞艳	刘卫英	吕 洁	李长江
	刘亚斌	李中书	丁新淼	刘亚民
主 审	张绍周			

# 前言

在质量管理体系的建设中，数理统计方法是提高企业产品质量和经济效益非常重要的一项基本工具。在 ISO 9000 族标准中，对数理统计方法的应用提出了明确而严格的要求，将统计技术从质量管理的“一个要素”提升为“质量管理体系的基础”，强调“过程策划应能确保各过程按规定的方法和顺序在受控状态下进行”。

近年来，我国标准化管理部门积极采用国际标准，陆续制定了一系列关于数理统计技术的基础国家标准。在企业中贯彻实施这些标准，通过正确、科学的数理统计方法对产品生产过程自始至终进行必要的监测和控制，可以及时调整生产工艺，保证最终产品的质量，提高企业的经济效益。特别是在现代化的大生产企业中，统计技术的应用日益显得迫切而重要。

为了提高我国产品制造企业质量管理和质量统计人员的数理统计知识水平，编者依据国家标准，参阅了若干文献资料，结合我国产品制造产业的实际情况，编写了本书。为便于读者理解和应用，编写此书时，对统计学中繁杂的数学公式不作过多的介绍和推导，而将重点放在对产品生产过程中各环节质量控制指标的数理统计技术的具体应用上，以帮助质量管理人员尽快掌握一些最基本的数理统计技术和统计方法，并能尽快在生产实践中予以应用。

本书内容共分十章：①统计技术基本知识；②试验误差与数据处理；③从样本推断批产品质量示例；④过程控制中的统计技术；⑤质量管理统计工具；⑥显著性检验；⑦方差分析与试验设计；⑧回归分析；⑨统计抽样检验；⑩计算机在统计工作中的应用。读者可以根据工作需要，阅读其中有关章节。

第一章至第五章，是产品制造企业工程技术人员、质量检验人员、质量管理人员和质量统计人员应该掌握的基本数理统计知识、统计技术和统计方法。

第六章、第七章和第八章，可供工程技术人员、质量管理人员进行技术改造和研制新产品、新工艺、新方法时参阅。

统计抽样检验，是区别于以往按百分比抽样检验的一种抽样检验方法，正在我国各行业中推广应用。为更好地贯彻实施近年来颁布的一系列有关国家标准，掌握各项标准的实施方法，本书第九章对这方面的内容进行了归纳和简要介绍。

计算机的广泛应用，使以前统计过程中诸如标准偏差、方差分析、回归分析等的复杂计算可以非常便捷地完成，大大节省了统计工作的步骤和时间。本书第十章对计算机在统计工作中的一些应用方法进行了介绍。

本书层次分明，图文并茂，列有大量实例。产品制造行业门类繁多，有的实例虽限于某一行业，但其基本原理和过程是相通的，各行业的统计人员可以举一反三，将所举实例转化到自己行业的统计工作中。

本书适用于产品制造企业工程技术人员、质量检验人员、质量管理人员和质量统计人员阅读使用，也可供各级产品质量监督检验机构技术人员参考，是适用于各种产品制造企业检验人员和质量管理、统计人员的通用参考书籍。高校学生学习数理统计课程时也可参阅。

限于编者的水平，本书所述内容难免存在疏漏和不当之处，敬请读者不吝指正。

编 者  
2014年1月于北京

# 目 录

## 第一章 统计技术基本知识 ..... 1

- 第一节 统计技术和质量管理 / 1
  - 一、统计技术的创立和发展 / 1
  - 二、统计技术在产品生产质量管理中的应用及本书内容的安排 / 1
- 第二节 统计技术基本概念 / 3
  - 一、统计数据 / 3
  - 二、统计技术、统计方法和统计工具 / 4
- 第三节 总体和样本 / 4
  - 一、总体 / 4
  - 二、样本 / 4
  - 三、样本分布的特征值 / 5
- 第四节 随机变量的分布 / 12
  - 一、随机变量 / 12
  - 二、分布的概念 / 12
  - 三、计量值数据的正态分布 / 12
  - 四、几种特征值的分布 / 16
- 第五节 有效数字与计算法则 / 17
  - 一、有效数字的位数 / 17
  - 二、近似数的运算规则 / 19
  - 三、数值修约规则 / 20
- 第六节 实验室样品的采取 / 21
  - 一、在不同地点采取样品的方法 / 22
  - 二、随机取样和系统取样 / 24

## 第二章 试验误差与数据处理 ..... 27

- 第一节 误差及其表示方法 / 27
  - 一、误差与偏差 / 27
  - 二、误差类型与产生误差的原因 / 27
  - 三、误差的表示方法 / 28
  - 四、误差的合成 / 29
  - 五、误差的正态分布 / 29
  - 六、准确度与精密度 / 30
- 第二节 正态样本离群值的判断和处理 / 32
  - 一、几种简单的检验方法 / 32
  - 二、GB/T 4883—2008 推荐的方法 / 34
- 第三节 试验数据的正态检验 / 40
- 第四节 测试结果的报告 / 42
  - 一、重复性限和再现性限 / 42

- 二、平行测试结果精密度的检查方法 / 43
- 三、多次平行测定结果的报出 / 45

### 第三章 从样本推断批产品质量示例..... 47

#### 第一节 水泥生产过程质量控制指标合格率的计算 / 47

- 一、水泥生料质量合格率的计算 / 47
- 二、水泥熟料质量合格率的计算 / 48
- 三、水泥质量指标的计算 / 51

#### 第二节 从正态分布参数推断质量合格率 / 58

- 一、正态分布特征参数与产品质量的关系 / 58
- 二、出磨生料质量合格率与标准差的关系 / 64
- 三、出窑熟料质量合格率与标准差的关系 / 70
- 四、出厂水泥质量合格率与标准差的关系 / 74

### 第四章 过程控制中的统计技术..... 77

#### 第一节 质量波动的原因 / 77

#### 第二节 工序能力与工序能力指数 / 78

- 一、工序能力 / 78
- 二、工序能力指数 / 78
- 三、工序能力分析 / 82
- 四、用工序能力指数计算水泥质量指标 / 83

#### 第三节 统计过程控制图 / 84

- 一、控制图原理 / 84
- 二、控制图的判断 / 86
- 三、应用控制图前应当考虑的问题 / 88

#### 第四节 常用计量值控制图 / 91

- 一、平均值和极差控制图 ( $\bar{x}$ - $R$  图) / 91
- 二、单值和移动极差控制图 ( $x$ - $R_s$  图) / 94
- 三、中位数和极差控制图 ( $\tilde{x}$ - $R$  图) / 96
- 四、均值与标准差控制图 ( $\bar{x}$ - $s$  控制图) / 98

#### 第五节 预控图 (彩虹图) / 100

- 一、预控法的提出 / 100
- 二、预控法的区域划分 / 101
- 三、预控法的应用 / 101
- 四、预控图的统计原理 / 102
- 五、预控图对过程控制的弹性管理 / 103
- 六、预控法的优点 / 103
- 七、应用预控图时应注意的问题 / 105

### 第五章 质量管理统计工具..... 106

#### 第一节 调查表 / 106



第二节	排列图 / 108
第三节	分层法 / 110
第四节	直方图 / 111
	一、直方图的定义和应用 / 111
	二、正态概率纸的应用 / 115
第五节	饼分图 / 119
	一、图形和用途 / 119
	二、画法 / 119
	三、注意事项 / 119
第六节	散布图 / 120
	一、散布图的定义和应用范围 / 120
	二、散布图的应用程序 / 120
第七节	非数字数据统计方法 / 122
	一、因果图、系统图和关联图 / 123
	二、PDPC法 / 125
	三、KJ法 / 125
	四、矩阵图 / 126
	五、矢线图(网络计划) / 126
	六、流程图 / 126

## 第六章 显著性检验 ..... 129

第一节	显著性检验的基本原理和具体步骤 / 129
	一、统计推断过程 / 129
	二、显著性检验的依据 / 130
	三、显著性检验的两类错误 / 131
	四、双侧显著性检验与单侧显著性检验 / 132
	五、显著性检验的主要步骤 / 133
第二节	正态分布总体分布中心 $\mu$ 的显著性检验 / 134
	一、 $u$ 检验法——总体标准差 $\sigma$ 已知时 / 134
	二、 $t$ 检验法——总体标准差 $\sigma$ 未知时 / 138
第三节	正态分布总体方差 $\sigma^2$ 的显著性检验 / 143
	一、一个正态总体的情形—— $\chi^2$ 检验法 / 143
	二、两个正态总体的情形—— $F$ 检验法 / 147
第四节	显著性检验的次序 / 150
	一、首先采用 $\chi^2$ 检验法进行检验 / 150
	二、首先采用 $F$ 检验法进行检验 / 151
第五节	非正态分布总体的显著性检验 / 154
	一、符号检验法 / 154
	二、秩和检验法 / 155

## 第七章 方差分析与试验设计 ..... 157

第一节	方差分析 / 157
-----	------------

- 一、基本概念 / 157
- 二、单因素方差分析 / 158
- 三、双因素方差分析 / 162
- 第二节 试验设计 / 164
  - 一、正交设计的基本方法 / 164
  - 二、正交设计的方差分析 / 168

## 第八章 回归分析 ..... 171

- 第一节 一元线性回归方程 / 171
  - 一、一元线性回归方程的建立 / 171
  - 二、一元线性回归方程的显著性检验 / 174
  - 三、预测值与实测值的比较 / 177
  - 四、利用回归方程对生产过程进行控制 / 178
- 第二节 一元非线性回归方程 / 178
  - 一、化非线性回归为线性回归 / 178
  - 二、常见曲线类型及变换公式 / 181
- 第三节 二元线性回归方程 / 183
  - 一、求方程参数 / 183
  - 二、求相关系数  $r$  / 184
  - 三、求回归方程的精度 / 184
  - 四、回归系数  $b_1$ 、 $b_2$  的比较判别 / 184

## 第九章 统计抽样检验 ..... 188

- 第一节 统计抽样检验概述 / 188
  - 一、抽样检验的基本概念 / 188
  - 二、统计抽样检验的术语和定义 / 189
  - 三、统计抽样检验的类型 / 192
- 第二节 周期检验标准 / 194
  - 一、周期检验的含义 / 194
  - 二、GB/T 2829—2002 的使用程序 / 194
  - 三、周期检验后的处置 / 197
- 第三节 验收统计抽样检验标准 / 198
  - 一、计数型验收统计抽样检验标准 / 198
  - 二、计量型验收统计抽样检验标准 / 201
- 第四节 质量监督检验统计抽样检验标准 / 210
  - 一、计数型统计抽样检验标准 GB/T 2828.4—2008 / 210
  - 二、计数型统计抽样检验标准 GB/T 2828.11—2008 / 213
  - 三、计量型统计抽样检验标准 GB/T 6378.4—2008 / 214
- 第五节 统计抽样检验的后续工作 / 218
  - 一、复查 / 218
  - 二、批的再提交和不合格品处理 / 218

**第十章 计算机在统计工作中的应用** ..... 220

- 第一节 计算一组数据的参数 / 220
  - 一、计算一组数据的平均值  $\bar{x}$  / 220
  - 二、求一组数据的最大值 / 221
  - 三、求一组数据的最小值 / 221
  - 四、求一组数据的中位数  $\tilde{x}$  / 221
  - 五、求一组数据的平均绝对偏差  $\bar{d}$  / 221
  - 六、求一组数据的众数 / 221
  - 七、求一组数据的偏差平方和 / 221
  - 八、求一组数据的方差  $s^2$  / 222
  - 九、求一组数据的实验标准差  $s$  / 222
- 第二节 计算一元线性回归方程的参数 / 222
  - 一、求一元线性回归方程的斜率  $b$  / 223
  - 二、求一元线性回归方程的截距  $a$  / 223
  - 三、求一元线性回归方程的相关系数  $r$  / 223
  - 四、求回归后对  $y$  的预期值 / 224
- 第三节 方差分析 / 224
  - 一、单因素方差分析的计算 / 224
  - 二、双因素方差分析的计算 / 226

**附录** ..... 228

- 附录一 有关数表 / 228
  - 附表 1 正态分布的双侧位数 ( $u_\alpha$ ) 表 / 228
  - 附表 2-1 正态分布表 1 / 228
  - 附表 2-2 正态分布表 2 / 230
  - 附表 3  $t$  检验临界值表 / 231
  - 附表 4  $\chi^2$  检验临界值表 / 232
  - 附表 5  $F$  检验临界值表 1 / 233
  - 附表 6  $F$  检验临界值表 2 / 238
  - 附表 7 符号检验表 / 239
  - 附表 8 秩和检验表 / 240
  - 附表 9 随机数表 / 241
- 附录二 《水泥企业质量管理规程》 / 243
- 附录三 《水泥取样方法》(GB/T 12573—2008) / 247
- 附录四 《评定水泥强度匀质性试验方法》(JC/T 578—2009) / 251
- 附录五 《通用水泥质量等级》(JC/T 452—2009) / 255

**参考文献** ..... 257

# 第一章 统计技术基本知识

## 第一节 统计技术和质量管理

### 一、统计技术的创立和发展

客观世界是十分复杂的，事物是互相联系、互相依存的，同时也是不断变化、不断运动、不断革新、不断发展的。人们在认识和改造客观世界的实践中，必须研究事物的数量关系，做到心中有数。对数量关系的研究，开始人们只知道数量的四则运算和其他简单的代数运算，这是常量间的数量关系。后来数学家创立了微分学，人们开始可以研究变量之间的数量关系。自变量与因变量之间有完全确定的关系，称之为函数关系。在客观世界中还有一类数量关系，变量之间虽然也存在依存关系，但并不完全确定，它们之间只存在某种大致的数量关系，这就是相关关系。客观世界中，具有这种相关关系的事物是随处可见的。实际上，许多存在函数关系的变量，由于试验或测定的误差，即使对同一自变量的值，其因变量也常常呈现微小的波动。这是因为有许多偶然的微小因素在起作用。有时，微小的偶然因素还能给人以假象。人们为了正确地认识世界，必然要从偶然的因素中辨认出主要的系统因素。这就是统计技术所要解决的问题。

反过来说，统计技术也是人们在长期不断地从偶然现象中揭示客观规律而逐渐形成的。统计技术的创立和发展是与数学的另一分支——概率论密切联系在一起的。自然科学发展到一定的高度，生产进入到机械化大生产时代，由于需要解决大量的技术问题，推动了早期概率论的发展，使它逐步成为一门独立的数学分支，体现了其真正的价值。20世纪30年代以来，工业生产高速化、自动化，更加促进了作为概率论的应用科学——数理统计的迅速发展。另一方面，科学试验在这个时期也成为人们认识世界的重要手段，与以往就生产管理方面所做的简单试验不同，人们很自然地要求提高试验效率，用较少的试验次数获得较多的试验信息，于是就发展了试验设计的理论与方法。目前，统计技术已在科学和生产的各个方面得到广泛的应用，特别是在质量管理方面的应用更为引人注目。

### 二、统计技术在产品生产质量管理中的应用及本书内容的安排

在产品生产的各道工序都可看到原材料、中间产品及最终产品质量变异的存在。统计技术有助于对变异进行测量、表述、分析、解释和建模，甚至使用相对有限的也能做到这一点。而对数据进行统计分析则有助于更好地理解变异的性质、程度和原因，从而有助于解决甚至预防由这些变异所可能引发的问题。并且，统计技术能使企业更好地利用可获得的数据做出决策，因而有助于企业在质量管理各过程包括产品实现过程中改进产品、过程的质

量,持续提高质量管理体系的有效性。下述内容就是统计技术在产品生产质量管理中的一些主要应用。

(1) 对产品原材料、中间产品及最终产品的质量进行估计 产品生产所用原材料、中间产品及最终产品的量是非常大的,对其化学成分和物理性质不可能全部加以测定,只能是测定总体中的一部分,即样本。数理统计技术就是指导人们如何通过样本的质量对总体的质量进行推断。如对出磨生料质量的控制,是通过每小时取一个样品测定氧化钙的含量,每班生料中氧化钙的含量即可由8个样品的测定值的平均值进行推断,然后与控制值进行比较,以确定该班生料氧化钙的含量是否合格。本书第三章主要介绍如何应用统计技术从样本的测定结果推断中间产品以及最终产品质量的方法。

(2) 预测生产过程中出现的不合格趋势,以便采取有效的措施加以预防 在产品生产过程中不合格品的出现一般都会有先兆。为了使在岗人员能及时发现潜在的不合格因素,统计技术可以提供一套科学的方法,就是从生产过程中取一定容量的样本进行观测分析,通过计算画出质量控制图,将这种质量控制图贴在车间里,再将日常生产检验数据点在图上。如果点靠近警戒线或上下限定线,就显示快出不合格品了,必须立即采取措施克服这种不良苗头。目前这种质量控制图已广泛地应用在生产第一线,成为保证生产处于受控状态、保证产品质量的有力工具。本书第四章主要介绍产品生产过程中各种质量控制图的使用方法。

(3) 分析影响生产过程的主要因素 产品生产中各道工序互相之间是紧密联系的,下一道工序的产品受到前一道工序许多因素的影响。通常这些因素分为两类,一类是系统因素,另一类是随机因素。人们要知道的是系统因素是否存在以及它的作用大小,但由于随机因素的干扰,往往分不清什么是系统因素,什么是随机因素。数理统计中的方差分析能指导人们如何找出影响显著的系统因素以及影响不显著的随机因素,并估计它们的大小。本书第五章介绍分析影响生产过程的各种因素的统计方法。

(4) 比较两个数量之间的真正差异 在产品生产过程中常需对两个不同的量进行比较。例如,按某种生料配比生产出来的水泥熟料的强度与另一种生料配比生产出来的熟料强度是否有显著性差异。为了进行这种比较,需要取若干样品进行测定。人们真正需要比较的是两者总体的性质,但因为不可能对所有熟料都进行测定,两批熟料总体的强度平均值都是不可知的。人们能做到的只是从两批熟料中各抽取一部分熟料进行测量。数理统计的显著性检验可通过这两个样本数据的比较,对总体的性质做出正确的结论,从而使人们不会因为个别的现象、个别的数据做出有关总体的错误结论。本书第六章主要介绍数理统计的显著性检验方法。

(5) 试验设计与方差分析 进行科学试验是任何工业取得技术进步不可缺少的手段。采用合理的试验设计,可用较少的试验代价获得较多的试验信息。同时因为有了比较科学的试验设计,还可寻求各因素最优的组合以获得最佳的工艺效果,结合方差分析还可找出各因素之间的相互关系。本书第七章介绍数理统计中方差分析与试验设计的方法。

(6) 建立因变量与自变量之间的回归关系 在事物的许多相关变量中,一类变量是可以调整和控制的,与之相关的另一类变量则随之变化,而且还受随机因素的影响而有微小的波动。在数理统计中,一个变量随着另一个或多个自变量的变化而变化的定量关系称为回归关系。例如水泥的28d抗压强度随3d抗压强度而变化。统计技术可以提供一套完整的回归分析方法,建立28d抗压强度对3d抗压强度的依存关系式,即回归方程。有了这种回归方程,

即可以用3d强度来对28d强度进行预测或控制。本书第八章主要介绍数理统计中的回归分析方法。

(7) 产品的验收和质量考核 上一道工序生产出来的中间产品,经检验部门检验合格后要验收入库,转给下一道工序,以保证最终的出厂产品质量全部合格;质检部门要定期对企业产品的质量进行考核;原材料采购部门购进一批原料后也要验收入库,因此产品的验收和考核是产品生产中的重要的一环。对于一批产品的验收,一般的做法是从产品中抽取一定数量的样品进行检验,根据检验的结果及规定的标准确定接收还是拒收这批产品。由于产品生产的批量十分大,而且物理、化学性能的测定又是破坏性的,只能抽取少量样品进行检验,用样本的检验结果对总体进行推断。既然是推断就有推断的误差。一个简单的抽样验收方案,包括样本大小(所需抽取样品进行检验的数量)和接受界限。这种抽样方案既可减少合格批的产品因推断误差而被拒收,也可减少不合格批的产品因推断误差而被接收,从而减少生产方和使用方的风险。本书第九章介绍统计抽样检验方法。

(8) 计算机在产品质量控制中的应用 现在开发的一些计算机应用程序使得以前让人望而生畏的复杂计算变得十分简单快捷,这就为统计技术在产品生产控制中的应用提供了极为实用而有力的工具。本书第十章介绍统计技术中一些常用的计算机程序的使用方法。

## 第二节 统计技术基本概念

统计技术是以概率论为基础的应用数学的一个分支。统计技术是研究随机现象中存在的统计规律的学科,是指应用有关的统计方法,收集、整理、分析和解释统计数据,并对其所反映的问题的性质、程度和原因做出一定结论的科学技术。

统计技术包括统计推断和统计控制两大内容。统计推断是指通过对样本数据的统计计算和分析,提供表示事物特征的数据,比较两个事物之间的差异,分析影响事物变化的原因,找出产品形成全过程中质量变化的规律,对总体质量水平进行推断,预测尚未发生的事件;统计控制是指通过对样本数据的统计计算和分析,采取措施消除过程中的异常因素,以保证产品质量特性的分布基本保持在设定值附近,使生产过程达到稳定受控状态。

### 一、 统计数据

数据是统计的对象。习惯上把由数字组成的数字数据称为数据。

(1) 数字数据 数字数据指由数字(0、1、2、3、4、5、6、7、8、9)和小数点组成的数据。数字数据是对可定量描述的特性的表达,可以通过抽样、测量、记录获得。任何数字数据又都可以形成(服从)一定的分布(统计规律)。

(2) 数据的分类

① 计量值数据。计量值数据是指可以连续取值,在有限的区间内可以无限取值的数据。长度、面积、体积、质量、密度、电压、电流、强度等,大部分质量特性的数值都属于计量值数据。

② 计数值数据。计数值数据是只能间断取值,在有限的区间内只能取有限数值的数据,如到会的人数、某天生产的产品件数、水泥安定性不合格品数、产品表面的缺陷数等。计数值数据是以正整数(自然数)的方式表现的。计数值数据又分为计件值数据和计点值数据。

## 二、统计技术、统计方法和统计工具

数理统计中常使用三个名词：统计技术、统计方法和统计工具。这三种提法有其共性，即均是研究随机现象中确定的数字规律，但也有其各自的特点。

(1) 统计技术 统计技术是一个大的概念，是就整个学科而言的，指的是一门技术的总概括。

(2) 统计方法 统计方法是指统计技术中的具体方法，如控制图、直方图、散布图等各是统计技术中的一种方法。原则上应称控制图、直方图、散布图等为统计方法。

(3) 统计工具 统计工具指简化的统计方法。统计技术的理论基础是概率论，但对这一理论，初级技术人员难以掌握，因此妨碍了统计技术的推广应用。为此，针对基层工人和初级技术人员的特点，质量管理专家开发了因果图、排列图、调查表、直方图、散布图、控制图和分层法，称为质量管理七种工具。随着质量管理的不断深化，20世纪70年代质量管理专家又开发出了系统图、关联图、矩阵图、矢线图、KJ法、PDPC法和矩阵数据解析法，称为质量管理新七种工具。所谓工具，指不讲统计方法的原理和设计，也不讲对统计结果的分析，只讲操作步骤。

## 第三节 ▶▶ 总体和样本

### 一、总体

研究或统计分析对象的全体元素组成的集合称为总体。总体具有完整性的内涵，是由某一相同性质的许多个别单位（元素或个体）组成的集合体。当总体内所含个体个数有限时，称为有限总体；当总体内所含个体个数无限时，称为无限总体。在统计工作中，可以根据产品的质量规程或实际工作需要，选定总体的范围，例如每个月出厂的建材产品、某一批进厂煤或原材料，都可视为一个总体。

总体的性质取决于其中各个个体的性质，要了解总体的性质，理论上必须对全部个体的性质进行测定，但在实际中往往是不可能的。一是在多数情况下总体中的个体数目特别多，可以说接近于无穷多，例如出厂的水泥，即使按袋计数，也不可能对所有的袋进行测定；二是组成总体的个体数是无限的，例如对一种新分析方法的评价分析，每次测定结果即为一个个体，可以一直测定下去永无终止；三是有些产品质量的检测是破坏性的，不允许对其全部总体都进行检测。基于总体的这种种情况，在实际工作中只能从总体中抽取一定数量的、有代表性的个体组成样本，通过对样本的测量求出其分布中心和标准差，借助于数理统计手段，对总体的分布中心 $\mu$ 和标准差 $\sigma$ 进行推断，从而掌握总体的性质。

### 二、样本

来自总体的部分个体的集合，称为样本。从总体获得样本的过程称为抽样。样本中的每个个体称为样品。样本中所含样品的个数，称为样本容量或样本大小。若样本容量适当地大，并且抽样的代表性强，则通过样本检测得到的分布特征值，就能很好地代表总体的分布特征值。总体和样本的关系如图1-1所示。

例如，在水泥生料配制过程中，为控制生料的质量，每小时从生料生产线上采取一个样

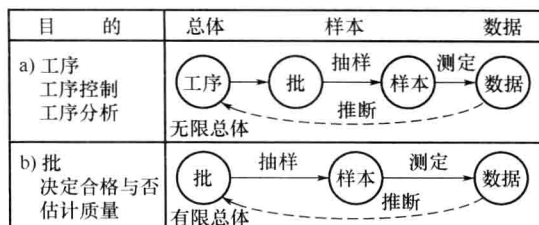


图 1-1 总体和样本的关系

品，进行硅、铁、铝、钙含量的测定。每天共采取 24 个样品，构成该日配制的生料总体的一个样本。对该样本中的 24 个样品的化学成分进行测定，可计算出该日配制的生料三率值的平均值。还可推广到整个生料库，将该生料库容纳的全部生料作为一个总体，其中每小时采取的样品之和作为样本，根据样本中所有样品的分析结果，计算该生料库中全部生料的三率值。

### 三、 样本分布的特征值

总体的分布特征值一般是很难得到的，数理统计中往往通过样本的分布特征值来推断总体的分布特征值。在实际应用中，为了对总体情况有一个概括的全面了解，需要用几个数字表达出总体的情况。这少数几个数字在数理统计中称为特征值。因此，在进行统计推断前确定样本分布的特征值，具有重要的实用价值。

常用的样本分布特征值分为两类：一类是位置特征值；另一类是离散特征值。

位置特征值一般是指平均值，它是分析计量数据的基本指标。在测量中所获得的检测数据都是分散的，必须通过平均值将它们集中起来，反映其共同趋向的平均水平，也就是说平均值表达了数据的集中位置，所以对一组测定值而言，平均值具有代表性和典型性。位置特征值一般包括算术平均值、几何平均值、加权平均值、中位数、众数等。

离散特征值用以表示一组测定数据波动程度或离散性质，是表示一组测定值中各测定值相对于某一确定的数值而言的偏差程度。一般是把各测定值相对于平均值的差异作为出发点进行分析。常用的离散特征值有平均差、极差、方差、标准（偏）差、变异系数等。

#### 1. 表示样本分布位置的特征值（样本分布中心）

(1) 算术平均值  $\bar{x}$  算术平均值的计算十分简单。它利用了全部数据的信息，具有优良的数学性质，是实际中应用最为广泛的反映样本集中趋势的度量值。

将一组测定值相加和，除以该组样本的容量（测定所得到的测定数据的个数），所得的商即为算术平均值。设有一组测定数据，以  $x_1, x_2, \dots, x_n$  表示。这组数据共由  $n$  个数据组成，其算术平均值为：

$$\bar{x} = (x_1 + x_2 + \dots + x_n) / n$$

或表示为：

$$\bar{x} = \sum_{i=1}^n x_i / n \quad (1-1)$$

式中  $n$ ——样本的容量；

$\sum_{i=1}^n$ ——在数理统计中，大写希腊字母  $\Sigma$  表示加和； $\Sigma$  下方的  $i=1$ ，表示从第一个数据



开始加和，一直加和到 $\Sigma$ 上方所表示的第 $n$ 个数据（在所指明确时，为了简化，有时只用 $\Sigma$ 表示，而不注出下方的 $i=1$ 和上方的 $n$ ）。

(2) 加权平均值 加权平均值是考虑了每个测量值的相应权的算术平均值。将各测量值乘以与其相应的权，将各乘积相加后，除以权数之和，即为加权平均值。其计算公式如下：

$$\bar{x}_w = \frac{p_1 x_1 + p_2 x_2 + \cdots + p_n x_n}{p_1 + p_2 + \cdots + p_n} = \frac{\sum p_i x_i}{\sum p_i} \quad (1-2)$$

式中  $x_1, x_2, \cdots, x_n$ ——各测量值；

$\bar{x}_w$ ——加权平均值；

$p_1, p_2, \cdots, p_n$ ——各测量值相应的权；

$\sum p_i$ ——各相应权的总和；

$\sum p_i x_i$ ——各测量值与相应权乘积之和。

根据实际情况，加权平均值有以下几种计算方法。

① 数量上的加权平均值。

**【例 1-1】** 水泥企业计算某一时期内熟料的综合抗压强度时，应采用加权平均值。某水泥厂有三台煅烧窑。其中 1 号窑年产 50 万吨熟料，抗压强度为 58.5MPa；2 号窑年产 60 万吨熟料，抗压强度为 57.8MPa；3 号窑年产 80 万吨熟料，抗压强度为 59.2MPa。

**【解】** 该厂全年水泥熟料综合抗压强度的加权平均值为：

$$(50 \times 58.5 + 60 \times 57.8 + 80 \times 59.2) / (50 + 60 + 80) = 11129 / 190 = 58.57(\text{MPa})$$

② “重要程度”的加权平均值。对于重要程度较高的特性，可以赋予较高的权；对于一般的特性，可以赋予较低的权。

③ 不同精度的加权平均值。不等精度测量时，由于获得各个测量结果的条件有所不同，各个测量结果的可靠性不一样，因而不能简单地取各测量结果的算术平均值作为最后的测量结果，应让可靠程度高的测量结果在最后的测量结果中占的比率大一些，可靠性低的占的比率小一些。

权的大小取决于测量值的可靠程度，而测量值的可靠程度又取决于它们各自的方差。方差愈小，可靠程度越高，权也就越大。因此，在不等精度测量列  $x_1, x_2, \cdots, x_m$  中，测量值的权与其相对应的方差成反比，即：

$$p_1 : p_2 : \cdots : p_m = \frac{1}{\sigma_1^2} : \frac{1}{\sigma_2^2} : \cdots : \frac{1}{\sigma_m^2} \quad (1-3)$$

式中  $p_i$ ——各测量值的权；

$\sigma_i$ ——各测量值的标准差， $i=1, 2, \cdots, m$ 。

**【例 1-2】** 有两个实验室用精度不同的量具测量同一物体的厚度，分别得到下述结果。欲得出其平均值，采用加权平均的办法。

**【解】**  $\bar{x}_1 = 1.53\text{mm}$ ， $\sigma_1 = 0.06\text{mm}$

$$\bar{x}_2 = 1.47\text{mm}$$
， $\sigma_2 = 0.02\text{mm}$

显然，第二个实验室的测定精度较高，其权数亦应大。两次测量的权数分别为：

$$p_1 = 1 / (0.06)^2 = 300$$

$$p_2 = 1 / (0.02)^2 = 2500$$